JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 6, pp. 1659-1668, Jun. 2026

Check for updates

실시간 영상 기반 도시홍수 침수영역 탐지 경량 세그멘테이션 AI 모델 개발

김 민 석[1] · 윤 천 주[2*]
[1]과학기술연합대학원대학교 한국건설기술연구원 스쿨 박사과정
[2]한국건설기술연구원 도로교통연구본부 수석연구원

Development of a Lightweight AI Segmentation Model for Detecting Flooded Areas in Real-Time Urban Flood Videos

Min Seok Kim[1] · Chunjoo Yoon[2*]

[1]Ph.D Course, KICT, University of Science and Technology, Goyang 10223, Korea
[2]Senior Researcher, Department of Highway and Transportation Research, KICT, Goyang 10223, Korea

[요 약]

본 연구는 도시홍수 CCTV 영상 기반 실시간 침수영역 세그멘테이션을 위해 10종의 AI 모델을 동일 조건에서 비교·평가하였다. 부산시 침수위험 복합 데이터셋 89,995장을 활용하여 mIoU, 클래스별 IoU, FPS, 파라미터 수를 분석한 결과, 일부 경량 모델은 높은 FPS를 보였으나 mIoU와 클래스별 IoU에서 성능 저하가 나타났다. 이에 본 연구에서는 제한된 연산 자원에서 도시홍수 모니터링이 가능하고 CCTV 탑재가 가능한 경량 세그멘테이션 모델 FloodSP-Net을 설계하였다. 제안 모델은 mIoU 0.8018, 556.4 FPS, 1.29M 파라미터 수를 기록하였다. 기존 경량 모델들과 mIoU, FPS, 파라미터 수를 비교한 결과, FloodSP-Net은 일정 수준의 정확도와 실시간성 및 경량성을 확인하였다. 따라서 FloodSP-Net은 CCTV 기반 도시홍수 실시간 모니터링 환경에서 적용 가능한 경량 세그멘테이션 모델로 판단된다.

[Abstract]

This study compared and evaluated ten segmentation AI models under identical conditions for real-time flooded-area segmentation in urban flood CCTV videos. Using 89,995 images from the Busan Flood Risk Complex Dataset, the models were analyzed in terms of mIoU, class-wise IoU, FPS, and the number of parameters. The results showed that certain lightweight models achieved high FPS but suffered from performance degradation in mIoU and class-wise IoU. Therefore, this study designed FloodSP-Net, a lightweight segmentation model that enables urban flood monitoring under limited computational resources and can be deployed in CCTV-based monitoring environments. The proposed model achieved an mIoU of 0.8018, 556.4 FPS, and 1.29M parameters, showing reasonable performance in terms of mIoU, FPS, and parameter size compared to existing lightweight models. These results indicate that FloodSP-Net can serve as a practical lightweight segmentation model for CCTV-based real-time urban flood monitoring.

색인어 : 도시홍수, 의미론적 분할, 경량 AI 모델, 실시간 모니터링, CCTV 영상
Keyword : Urban Flooding, Semantic Segmentation, Lightweight AI Model, Real-Time Monitoring, CCTV Video

http://dx.doi.org/10.9728/dcs.2026.27.6.1659

Received 15 May 2026; **Revised** 01 June 2026
Accepted 15 June 2026

***Corresponding Author; Chunjoo Yoon**

Tel: +82-31-910-0533
E-mail: cjyoon@kict.re.kr

Ⅰ. 서 론

최근 기후변화와 도시화의 영향으로 도시지역의 집중호우 및 내수침수 위험이 증가하고 있다. IPCC 제6차 평가보고서에서는 도시화가 도시 및 도시 하류 지역의 평균 강수와 극한 강수를 증가시킬 수 있으며, 이로 인한 유출 강도 증가는 도로, 지하공간, 건축물 침수로 이어질 수 있음을 제시하였다[1]. 국내에서도 행정안전부 재해연보는 호우, 태풍 등 자연재난으로 인한 피해 현황, 복구비, 인명피해 등을 지속적으로 집계하고 있으며, 이는 도시홍수 대응 및 재난관리 정책 수립의 기초자료로 활용되고 있다[2]. 특히 도심부 도로 침수는 차량 고립, 교통 혼잡, 긴급차량 접근 지연, 2차 사고 유발 등 도로교통 안전 전반에 직접적인 영향을 미치며, 도시홍수 상황에서는 주요 도로의 기능 저하와 우회도로의 혼잡 증가가 도시 교통망 전반으로 확산될 수 있다[3]. 따라서 침수영역의 공간적 확산 양상과 도로 주변 구조물의 상태를 신속하게 파악할 수 있는 실시간 모니터링 기술이 필요하다.

기존 도시홍수 모니터링은 주로 우량계, 수위계, 하천 수위 관측장비 등 계측 기반 자료에 의존해 왔다. 이러한 방식은 특정 지점의 수위나 강우량을 정량적으로 파악하는 데 유용하지만, 실제 도로 공간에서 침수영역이 어느 방향으로 확산되는지, 도로 경계부와 배수시설 주변의 위험 상태가 어떻게 변화하는지를 공간 단위로 파악하는 데에는 한계가 있다. 반면 도시 전역에 설치된 CCTV는 현장의 시각 정보를 연속적으로 제공하므로, 도로 침수 상황을 직접 관측하고 침수영역과 주변 구조물을 동시에 분석할 수 있는 실용적 자료원으로 활용될 수 있다.

그러나 CCTV 영상을 도시홍수 모니터링에 활용하기 위해서는 침수영역과 도로 주변 구조물을 자동으로 탐지할 수 있는 AI 기반 세그멘테이션 기술이 필요하다. 도시홍수 CCTV 영상은 침수영역, 도로면, 교각, 배수시설 등이 복합적으로 나타나며, 이때 도로면은 주행공간으로의 침수 확산 여부를 파악하는 데 필요하며, 교각과 배수시설은 침수 확산 양상과 침수에 따른 위험 상태를 해석하는 데 필요한 요소이다. 따라서, 단순히 침수 여부를 판단하는 것보다 침수영역과 주요 구조물을 픽셀 단위로 구분하는 세그멘테이션 기술이 중요하다. 또한 현장 단말 장치와 같은 엣지 디바이스 기반 운영 환경을 고려하면 AI 모델은 탐지 정확도뿐만 아니라 실시간 추론 속도와 낮은 파라미터 수 확보가 필수적이라 판단된다.

이에 본 연구에서는 AI Hub의 부산시 침수위험 복합 데이터셋[4]을 활용하여 도시홍수 CCTV 영상의 데이터 구성과 클래스별 분포 특성을 분석하였다. 또한, 10종의 세그멘테이션 모델을 동일 조건에서 학습·평가하고, mIoU, 클래스별 IoU, FPS, 파라미터 수를 기준으로 도시홍수 CCTV 기반 엣지 환경에서의 적용 가능성을 분석하였다. 이를 통해 제한된 연산 자원에서 파라미터 수를 줄이면서 일정 수준의 mIoU와 실시간 추론 속도를 확보할 수 있는 경량 세그멘테이션 모델 FloodSP-Net을 제안하였다.

Ⅱ. 본 론

2-1 기존 문헌 고찰

도시홍수 CCTV 기반 침수영역 탐지와 관련된 선행연구는 크게 침수영역 세그멘테이션 연구와 실시간 경량 세그멘테이션 모델 연구로 구분할 수 있다.

먼저 침수영역 세그멘테이션 연구는 주로 위성영상, 항공영상, UAV 영상 등 상공 시점 자료를 중심으로 발전해 왔다. 허리케인 하비 이후 촬영된 고해상도 UAV 영상을 기반으로 구축된 FloodNet 데이터셋은 침수 도로, 침수 건물, 자연 수역 등을 구분하기 위한 post-flood scene understanding 문제를 제시하여, 고해상도 영상 기반 세그멘테이션의 재난지역 분석 활용 가능성을 보였다[5]. 또한 UAV 영상 기반 침수영역 매핑 연구에서는 VGG 기반 FCN-16s 모델과 k-fold 교차검증을 통해 침수영역 세그멘테이션 가능성을 검토하였다[6]. 그러나 이들 연구는 대부분 상공 시점 영상을 대상으로 하므로, 지상 CCTV 영상의 도로면, 교각, 배수시설 등 주요 구조물을 동시에 탐지하기에는 한계가 있다.

최근에는 CCTV 및 감시카메라와 같은 지상 시점 영상을 활용한 도시홍수 모니터링 연구가 확대되고 있다. Liang et al.은 V-FloodNet을 제안하여 이미지 및 비디오 기반 침수영역 세그멘테이션과 기준 객체 탐지를 수행하고, 이를 기반으로 도시홍수 수심을 추정하는 시스템을 제시하였다[7]. 해당 연구는 이미지 및 비디오 자료가 도시 공간에서 비교적 쉽게 확보될 수 있다는 점을 바탕으로, 영상 기반 도시홍수 모니터링의 활용 가능성을 보였다. Wang et al.은 실제 감시카메라 영상 기반 도시홍수 이미지에 DCNN 기반 세그멘테이션 모델을 적용하여 침수영역 추출 가능성을 검토하였으며, 반사 수면, 젖은 도로면, 저조도, 강우 중 카메라 빗방울 등이 침수영역 분할 성능에 영향을 미칠 수 있음을 제시하였다[8]. Zhao et al.은 저조도 감시영상 기반 도시 침수 모니터링 연구를 통해 야간 및 강우 상황에서 침수영역의 영상 특성이 복잡하게 변화하며, 이를 보완하기 위한 저조도 환경 특화 세그멘테이션 모델의 필요성을 제시하였다[9]. 이와 같이 최근 연구들은 지상 시점 영상이 도시홍수 모니터링에 활용될 수 있음을 보여주고 있으나, 도시홍수 CCTV 환경에서 침수영역과 도로 주변 구조물을 동시에 고려하면서 실시간 추론 속도와 경량성을 함께 분석한 연구는 여전히 제한적이다.

실시간 경량 세그멘테이션 모델 연구에서는 정확도와 추론 속도의 균형을 확보하기 위한 다양한 구조가 제안되어 왔다. DeepLabV3+는 Atrous Separable Convolution과 encoder-decoder 구조를 결합하여 다중 스케일 문맥 정보와 경계 복원 성능을 강화한 모델이다[10]. ENet은 낮은 지연시간과 적은 연산량을 목표로 한 초기 실시간 세그멘테이션 모델이다[11]. Fast-SCNN은 Learning to Downsample 모듈과 Feature Fusion 구조를 통해 저메모리 환경에서의 실시간 세그멘테이션을 목표로 설계되었다[12]. BiSeNetV2는

공간 정보와 의미 정보를 분리한 양측 분기 구조와 Guided Aggregation layer를 통해 정확도와 추론 속도의 균형을 확보하였다[13]. DDRNet은 고해상도·저해상도 분기를 병렬로 유지하는 dual-resolution 구조를 통해 도로 장면에서의 절충 성능을 개선하였다[14]. STDC는 Short-Term Dense Concatenate 구조와 Detail Aggregation 모듈을 통해 단일 스트림 기반 실시간 세그멘테이션 성능을 강화하였다[15]. PP-LiteSeg는 경량 디코더, Attention Fusion, Pyramid Pooling 구조를 결합하여 경량성과 정확도 간의 절충 성능을 제시하였다[16]. PIDNet은 Detail, Context, Boundary 정보를 처리하는 3분기 구조를 통해 경계 정보 보존 성능을 강화하였다[17]. SegFormer는 계층적 Transformer 인코더와 경량 MLP 디코더를 결합하여 효율적인 세그멘테이션 성능을 확보하였다[18].

기존 문헌은 침수영역 탐지, 수심 추정, 실시간 세그멘테이션 모델 개발 측면에서 연구 성과를 제시하였으나, 도시홍수 CCTV 영상에서 침수영역과 도로 주변 구조물을 동시에 고려하고, 경량 모델의 정확도·추론 속도·파라미터 수를 동일 조건에서 비교한 연구는 여전히 제한적이다.

2-2 기존 문헌 고찰의 시사점

기존 침수영역 세그멘테이션 연구는 주로 위성·항공·UAV 영상 등 원격탐사 기반 자료를 대상으로 수행되어 왔다. 이는 광역 침수영역 탐지에는 유용하지만, 침수영역과 주행공간의 관계, 배수시설 주변 침수 확산, 구조물 주변 위험 상태를 실시간으로 분석해야 하는 CCTV 기반 모니터링 환경과는 차이가 있다. 또한 CCTV 기반 실시간 모니터링은 다수 지점의 영상을 연속적으로 처리해야 하므로, 현장 단말 또는 엣지 디바이스의 연산 자원 제약을 고려해야 한다[7],[8]. 따라서 도시홍수 CCTV 영상에서는 침수영역과 주요 구조물을 동시에 탐지하면서도 빠른 추론 속도와 낮은 파라미터 수를 갖는 경량 세그멘테이션 모델이 필요하다. 이에 본 연구에서는 10종의 세그멘테이션 모델을 동일 조건에서 비교하고, 도시홍수 모니터링을 위해 CCTV에 탑재 가능한 경량 이중 분기 구조의 FloodSP-Net을 제안한다.

Ⅲ. 세그멘테이션 AI 모델 성능 분석

3-1 데이터셋 개요 및 분석

1) 데이터셋 개요

본 연구에서는 AI Hub에서 제공한 부산시 침수위험 복합 데이터셋을 활용하였다. 해당 데이터셋은 도시홍수 상황에서 촬영된 CCTV 영상을 기반으로 구축되었으며, 그림 1과 같이 세그멘테이션 학습을 위한 원본 이미지와 Ground Truth 마스크 이미지로 구성된다.

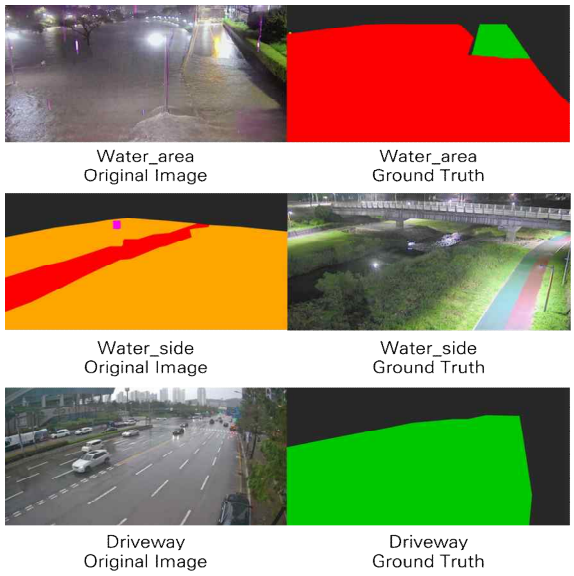

그림 1. 데이터셋 예시
Fig. 1. Example dataset

원본 영상은 실제 강우 이벤트 시 수집된 현장 데이터를 반영하고 있으며, 도시홍수 실시간 모니터링 환경에서 발생할 수 있는 다양한 시각적 조건과 침수 양상을 포함한 것을 알 수 있다. 데이터셋은 총 89,995장의 이미지로 구성되며, 이 중 79,995장은 학습용, 10,000장은 검증용으로 구성되어 제공된다. 모델 성능 평가는 해당 검증용 데이터를 기준으로 수행하였다. 원본 영상의 해상도는 1,280×720이며, 모델 간 공정한 비교와 학습 효율성 확보를 위하여 모든 입력 영상을 512×512 크기로 리사이즈하여 사용하였다. 데이터셋의 클래스는 Background, Driveway, Pole, Pier, Water_side, Water_area, Rain_gutter로 총 7개의 클래스로 구성된다.

2) 데이터셋 분석

그림 2와 같이 클래스별 픽셀 비율을 분석하면, Background 44.9%, Driveway 32.0%, Water_area 12.1%, Water_side 9.0%, Rain_gutter 1.24%, Pier 0.46%, Pole 0.19%로 나타난다. 즉, Background와 Driveway가 전체 픽셀의 76.9%를 차지하는 반면, Pole, Pier, Rain_gutter는 합계 1.89%에 불과하여 클래스 불균형 특성을 보인다.

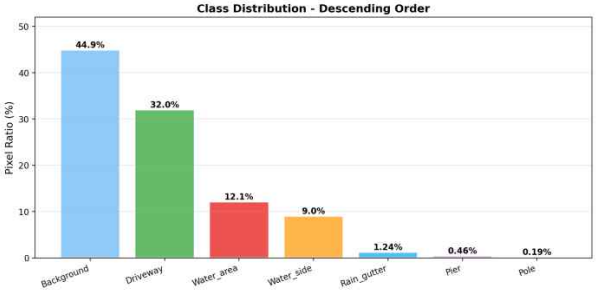

그림 2. 클래스별 픽셀 비율
Fig. 2. Pixel ratio per class

그림 3과 같이 Pole은 세로 방향으로 길고 얇은 형태를 보이며, Pier는 교각에 해당하는 세로 블록형 구조로 나타난다. Rain_gutter는 도로변 또는 배수시설을 따라 가로 방향으로 길게 분포하는 특성을 가진다. 해당 클래스들은 전체 픽셀에서 차지하는 비율이 낮고 방향성이 뚜렷한 소형·세장형 구조물로, 반복적인 다운샘플링 과정에서 공간 정보가 손실되기 쉽다. 따라서 도시홍수 CCTV 기반 세그멘테이션 모델은 넓게 분포하는 침수영역과 도로 영역을 안정적으로 구분하는 동시에, 상대적으로 작은 비율로 나타나는 구조물과 침수 경계부의 세부 정보를 보존할 수 있어야 한다. 이러한 데이터 특성은 소형 객체 보존, 소수 클래스 학습 보강, 방향성 구조물 인지를 고려한 FloodSP-Net 설계에 반영하였다.

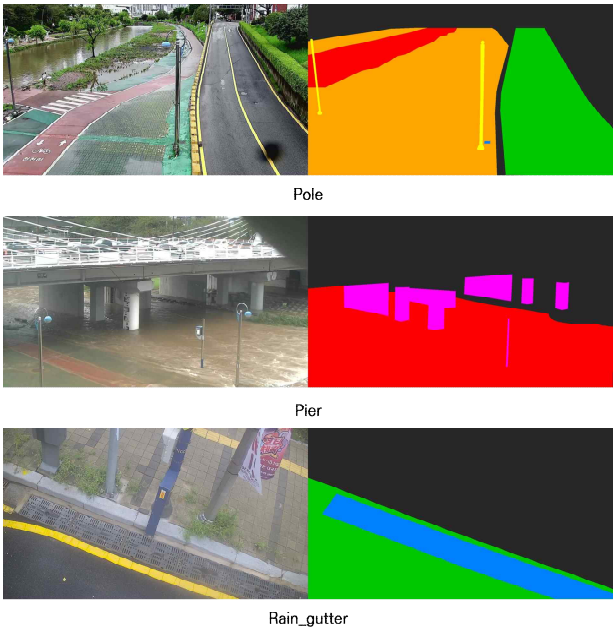

그림 3. Pole, Pier, Rain_gutter 클래스 원본 및 GT 이미지
Fig. 3. Original, GT images of Pole, Pier, Rain_gutter classes

3-2 비교 모델 특성 및 학습 조건

본 연구에서는 도시홍수 CCTV 영상 기반 실시간 침수영역 세그멘테이션에 적용 가능한 총 10종의 세그멘테이션 모델을 비교 대상으로 선정하였다. 비교 모델은 정확도 중심의 중량급 모델, 경량 백본 기반 모델, 이중 분기 구조 모델, 경량 실시간 모델을 포함하도록 구성하였다. 이는 도시홍수 CCTV 환경에서 요구되는 정확도, 추론 속도, 경량성 간의 상충 관계를 분석하고, 실제 모니터링 환경의 연산 자원과 처리 속도 요구조건을 고려하여 적정 모델을 도출하기 위함이다.

구체적으로 DeepLabV3+ (ResNet-50)는 고정확도 기준 모델로 선정하여 본 데이터셋에서 확보 가능한 성능 상한선을 확인하고자 하였다. DeepLabV3+ (MobileNetV3), STDC1-Seg, SegFormer-B0는 경량 백본 기반 모델로 포함하여, 경량화된 백본이 정확도, 추론 속도, 파라미터 수에 미치는 영향을 검토하고자 하였다. BiSeNetV2, DDRNet23-Slim, PIDNet-S는 공간 정보와 문맥 정보를 병렬, 이중 해상도 구조로 처리하는 경량 모델로서, 대면적 침수영역과 소형 구조물이 공존하는 도시홍수 장면에서의 적용 가능성을 평가하기 위해 선정하였다. ENet, Fast-SCNN, PP-LiteSeg는 경량 및 고속 추론 중심 모델로 분류하여, 제한된 연산 자원을 갖는 엣지 기반 실시간 모니터링 환경에서의 활용 가능성을 함께 분석하고자 하였다. 이와 같이 다양한 구조적 특성을 갖는 모델을 비교함으로써, 단순한 정확도 중심 평가가 아니라 도시홍수 CCTV 기반 침수영역 세그멘테이션에 적합한 모델의 성능 특성을 종합적으로 검토하고자 하였다. 선정한 10종 모델의 구성 및 특성은 표 1과 같다.

표 1. 10종 세그멘테이션 모델 구성
Table 1. Configuration of 10 segmentation models

No.	Model	Backbone	Category
1	DeepLabV3+	ResNet-50	Heavy baseline
2	DeepLabV3+	MobileNetV3-Large	Lightweight baseline
3	BiSeNetV2	Bilateral Detail-Semantic Network	Real-time CNN
4	PP-LiteSeg	STDC	Real-time CNN
5	PIDNet-S	PIDNet-S	Real-time CNN
6	SegFormer-B0	MiT-B0	Lightweight Transformer
7	DDRNet23-Slim	DDRNet-23	Dual-resolution CNN
8	STDC1-Seg	STDC1	Real-time CNN
9	ENet	ENet Encoder	Ultra-lightweight CNN
10	Fast-SCNN	Fast-SCNN	Edge-oriented CNN

모든 비교 모델은 공정한 성능 비교를 위하여 동일한 학습 조건에서 학습하였다. 입력 영상 크기는 512 × 512로 통일하였으며, 배치 크기는 16으로 설정하였다. 비교 대상 10개 모델의 학습 Epoch은 각 모델의 개별 최적 성능을 도출하기보다는 모델 간 비교 조건의 일관성과 GPU 자원 및 학습 시간 제약을 고려하여 15로 설정하였고, 옵티마이저는 AdamW를 사용하였다. 초기 학습률은 1e-4, Weight Decay는 1e-4로 설정하였으며, 학습률 스케줄러는 Cosine Annealing을 적용하였다. 손실함수는 클래스 불균형 특성을 반영하기 위하여 역빈도 기반 Weighted Cross-Entropy를 사용하였다. 데이터 증강은 모델 간 비교 조건의 일관성을 유지하기 위해 Random Horizontal Flip만 적용하였다. 이는 모든 비교 모델에 동일한 증강 조건을 적용하여 모델 구조에 따른 상대적 성능 차이를 비교하기 위한 설정이다. 특히 도시홍수 CCTV 영상은 도로, 침수영역, 구조물의 공간적 배치가 중요한 데이터이므로, 과도한 기하학적 변환이나 색상 변형은 클래스 경계 및 장면 구조를 왜곡할 수 있다. 따라서 본 연구에서는 기

본적인 좌우 반전 증강만 적용하였으며, 사전학습이 가능한 모델은 ImageNet pretrained 가중치를 활용하였다. 실험은 NVIDIA RTX PRO 4500 Blackwell (32 GB) 환경에서 수행하였다.

그림 4는 동일한 학습 조건에서 10종 세그멘테이션 모델의 학습 및 검증 Loss 변화를 나타낸다. 학습 Loss는 초기 1~3 Epoch에서 Loss가 급격히 감소하였고, 4~5 Epoch 이후부터는 완만하게 수렴하였다. 검증 Loss 또한 1~4 Epoch에서 빠르게 감소한 뒤, 5 Epoch 이후에는 안정화되었다. 이를 통해 15 Epoch 조건에서 모델들이 전반적으로 수렴하였음을 확인하였다.

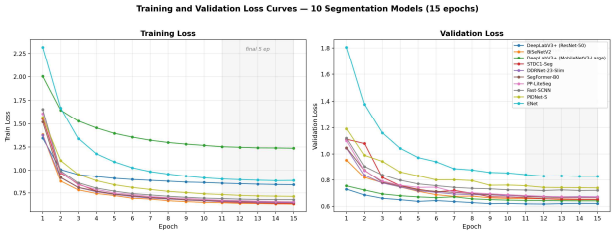

그림 4. 10종 세그멘테이션 모델의 학습 및 검증 loss 변화
Fig. 4. Training and validation loss curves of 10 segmentation models

3-3 모델 성능 비교

모델 성능은 mIoU, 클래스별 IoU, 추론 속도(FPS), 파라미터 수를 기준으로 평가하였다. mIoU는 전체 세그멘테이션 정확도를 나타내는 대표 지표로 사용하였으며, 클래스별 IoU는 도시홍수 CCTV 장면에서 클래스별 성능 차이를 분석하기 위해 활용하였다. 추론 속도는 실제 CCTV 모니터링 환경에서의 실시간 적용 가능성을 판단하기 위한 지표로, 배치 크기 1 기준 반복 추론을 통해 측정하였다. 파라미터 수는 경량 환경에서의 적용 가능성을 판단하기 위한 지표로 사용하였다. 본 연구에서 제시한 파라미터 수는 7개 클래스 세그멘테이션에 맞게 구현한 모델을 기준으로 산정하였다.

전체 모델의 정량적 성능 비교 결과는 표 2와 같다. DeepLabV3+ (ResNet-50)는 mIoU 0.9002로 가장 높은 세그멘테이션 성능을 기록하였다. 이는 도시홍수 CCTV 장면에서도 중량급 세그멘테이션 모델이 높은 정확도를 확보할 수 있음을 보여준다. 그러나 해당 모델은 41.99M의 파라미터 수를 요구하므로, 제한된 연산 자원을 갖는 엣지 환경에 직접 적용하기에는 한계가 있다. 반면 BiSeNetV2는 mIoU 0.8594, 추론 속도 600.9 FPS, 파라미터 수 2.04M을 기록하여 정확도와 효율성 측면에서 가장 우수한 균형을 보였다. 이는 실시간 도시홍수 CCTV 모니터링 환경에서 높은 세그멘테이션 성능과 경량성을 동시에 확보할 수 있는 대표적인 균형형 모델임을 시사한다.

표 2. 세그멘테이션 모델의 정량적 성능 비교
Table 2. Performance comparison of segmentation models

Model	mIoU	FPS	Parameters
DeepLabV3+ (ResNet-50)	0.9002	93.6	41.99M
BiSeNetV2	0.8594	600.9	2.04M
DeepLabV3+ (MobileNetV3-Large)	0.8411	455.8	11.03M
STDC1-Seg	0.8375	484.6	25.47M
DDRNet23-Slim	0.8274	617.0	3.68M
SegFormer-B0	0.8212	331.4	3.52M
PP-LiteSeg	0.8097	545.3	6.59M
Fast-SCNN	0.7914	784.5	1.12M
PIDNet-S	0.7560	1459.0	1.67M
ENet	0.6587	404.4	0.21M

경량 모델군인 Fast-SCNN, PIDNet-S, ENet은 모두 2M 이하의 파라미터를 확보하였다. 그러나 mIoU는 각각 0.7914, 0.7560, 0.6587로 나타나, 도시홍수 CCTV 장면의 복합적인 구조를 충분히 반영하는 데에는 한계가 확인되었다. 특히 ENet은 0.21M의 매우 작은 파라미터 수를 갖지만 정확도 저하가 크게 나타났으며, PIDNet-S는 가장 높은 FPS를 기록하였으나 mIoU가 낮았다. 해당 결과는 엣지 환경에서 단순히 파라미터 수를 줄이거나 추론 속도만을 높이는 것만으로는 도시홍수 CCTV 기반 세그멘테이션 성능을 충분히 확보하기 어렵다는 점을 보여준다. 따라서 본 연구에서는 mIoU, FPS, 파라미터 수를 종합적으로 고려하여 BiSeNetV2, Fast-SCNN, PIDNet-S, ENet을 주요 경량 비교 모델로 선정하고, 이를 FloodSP-Net과 비교하였다.

3-4 모델별 클래스별 성능 분석

mIoU는 전체 클래스의 평균 탐지 성능을 나타내는 지표이므로, 도시홍수 CCTV 복합 장면에서의 실질적인 모델 적용성을 충분히 평가하기 어렵다. 동일한 mIoU 수준의 모델이라도 특정 클래스에서 성능 차이가 발생할 수 있으므로, 본 연구에서는 10종 모델의 클래스별 IoU를 표 3과 같이 비교하였다.

분석 결과, Background와 Driveway와 같은 대면적 클래스에서는 BiSeNetV2와 경량 모델군 모두 비교적 안정적인 성능을 보였다. BiSeNetV2는 Background 0.9823, Driveway 0.9792를 기록하였으며, Fast-SCNN은 각각 0.9663, 0.9622, PIDNet-S는 0.9571, 0.9450, ENet은 0.9308, 0.9120으로 나타났다. 이는 대면적 클래스의 경우 경량 모델에서도 일정 수준 이상의 탐지 성능을 확보할 수 있음을 보여준다.

반면 Pole, Pier, Rain_gutter와 같은 소형·세장형 구조물 클래스에서는 모델 간 성능 차이가 뚜렷하게 나타났다. BiSeNetV2는 Pole 0.7724, Pier 0.7648, Rain_gutter 0.6277을 기록하여 경량 모델군 대비 가장 안정적인 성능을

표 3. 세그멘테이션 모델별 클래스 IoU 비교
Table 3. Class-wise IoU by segmentation models

Model	Background	Driveway	Pole	Pier	Water_side	Water_area	Rain_gutter
DeepLabV3+ (ResNet-50)	0.9929	0.9909	0.8233	0.8321	0.9701	0.9769	0.7150
BiSeNetV2	0.9823	0.9792	0.7724	0.7648	0.9365	0.9526	0.6277
DeepLabV3+ (MobileNetV3-Large)	0.9874	0.9839	0.6780	0.7262	0.9513	0.9652	0.5959
STDC1-Seg	0.9846	0.9772	0.7220	0.7474	0.9427	0.9550	0.5335
DDRNet23-Slim	0.9794	0.9782	0.6621	0.7057	0.9281	0.9500	0.5882
SegFormer-B0	0.9774	0.9752	0.6441	0.6921	0.9222	0.9440	0.5934
PP-LiteSeg	0.9839	0.9745	0.6642	0.6847	0.9368	0.9504	0.4734
Fast-SCNN	0.9663	0.9622	0.6423	0.6566	0.8876	0.9208	0.5038
PIDNet-S	0.9571	0.9450	0.5978	0.6559	0.8601	0.8666	0.4095
ENet	0.9308	0.9120	0.5054	0.4600	0.7707	0.7877	0.2443

보였다. 반면 Fast-SCNN은 각각 0.6423, 0.6566, 0.5038로 감소하였고, PIDNet-S는 0.5978, 0.6559, 0.4095를 기록하였다. ENet은 Pole 0.5054, Pier 0.4600, Rain_gutter 0.2443으로 가장 큰 성능 저하를 보였다. 특히 Rain_gutter 클래스는 모든 경량 모델에서 낮은 성능을 보였으며, 이는 가로 방향으로 길게 분포하고 픽셀 비율이 낮은 구조물의 특성이 경량 모델에서 충분히 보존되지 못했기 때문으로 해석된다.

이러한 성능 저하는 경량 모델의 반복적인 다운샘플링 구조와 소수 클래스의 데이터 불균형이 복합적으로 작용한 결과로 해석된다. 경량 모델은 연산량 감소를 위해 공간 해상도를 단계적으로 축소하므로, 픽셀 규모가 작은 객체의 세부 정보가 깊은 계층에서 손실되기 쉽다. 또한 정사각형 커널이나 Spatial Pooling 방식은 가로·세로 방향으로 길게 분포하는 구조물의 형태 정보를 충분히 반영하기 어렵고, 픽셀 비율이 낮은 소수 클래스는 학습 과정에서 충분한 그래디언트 신호를 확보하기 어렵다. 따라서 본 연구에서는 2M 이하의 경량 파라미터 규모에서 실시간성을 유지하면서, 대면적 침수영역과 소수 구조물 클래스의 특징 손실 문제를 고려한 경량 세그멘테이션 모델 FloodSP-Net을 설계하였다.

Ⅳ. FloodSP-Net 제안 및 성능 검증

4-1 FloodSP-Net 구조

FloodSP-Net은 Detail Branch, Context Branch, Multi-Scale Decoder로 구성된 경량 이중 분기 구조이다. 전체 파라미터 수는 1.29M이며, 입력 영상은 두 분기를 통해 각각 공간 세부 정보와 광역 문맥 정보를 추출한 뒤, Multi-Scale Decoder에서 단계적으로 융합된다. 모델의 전체 구조는 그림 5와 같다.

1) Detail Branch

Detail Branch는 입력 영상의 공간 정보를 보존하여 도로 윤곽, 침수 경계, 구조물의 형태 정보를 추출하는 역할을 수행한다. Depthwise Separable Convolution(DSConv) 블록을 기반으로 하며, 세 단계(D1, D2, D3)로 구성된다. 각 단계는 stride 2를 적용하여 입력 512×512×3으로부터 256×256×48, 128×128×64, 64×64×96의 특징맵을 순차적으로 생성한다. DSConv 블록은 3×3 depthwise convolution, batch normalization, ReLU, 1×1 pointwise convolution, batch normalization, ReLU의 순서로 구성된다. 이는 일반 합성곱 대비 연산량을 줄이면서도 공간 특징을 효과적으로 추출할 수 있는 구조이다. 따라서 Detail Branch는 경량성을 유지하면서도 도시홍수 장면에서 중요한 경계 및 세부 구조 정보를 보존하는 역할을 담당한다.

2) Context Branch

Context Branch는 광역 장면 문맥을 추출하기 위한 경로로, MobileNetV2 계열의 Inverted Residual(IR) 블록을 기반으로 한다. 일반적인 경량 모델이 1/32 해상도까지 다운샘플링하는 것과 달리, FloodSP-Net은 최저 해상도를 1/16으로 제한한다. 이는 도시홍수 장면에서 소형 구조물의 공간 특징이 과도하게 손실되는 것을 완화하기 위한 설계이다. 입력 영상은 stem 단계의 두 개의 convolution block을 통해 128×128×48 특징맵으로 변환된다. 이후 Stage C1은 IR 블록을 통해 64×64×64 특징맵을 생성하고, Stage C2는 32×32×160 특징맵을 생성한다. 마지막으로 Global Average Pooling(GAP)을 통해 전역 문맥 정보를 추출하고, 이를 32×32 특징맵에 보강하여 장면 수준의 의미 정보를 반영한다. Context Branch는 64×64 및 32×32 특징을 디코더로 전달하여, 단일 스케일 출력보다 풍부한 문맥 정보를 제공한다.

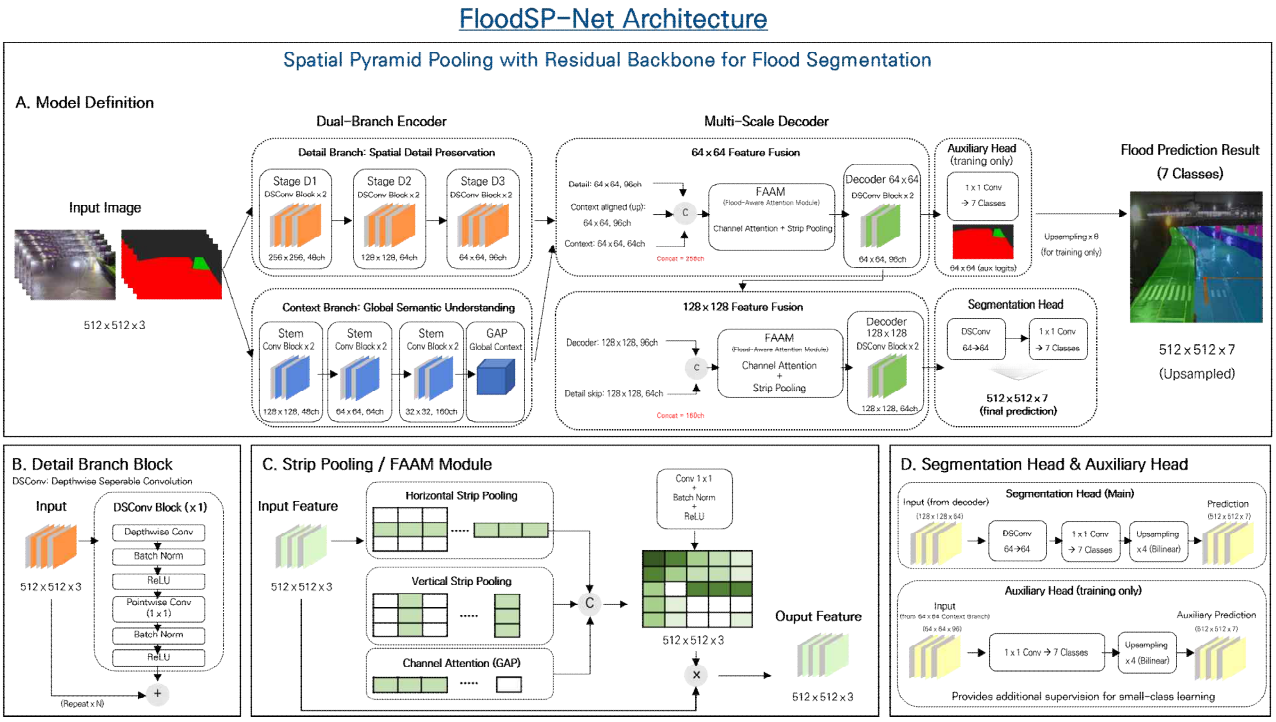

그림 5. FloodSP-Net 구조 및 특징

Fig. 5. FloodSP-Net architecture and characteristics

3) Strip Pooling 기반 Flood-Aware Attention Module

Flood-Aware Attention Module(FAAM)은 방향성 구조물의 형태 정보를 강화하기 위한 어텐션 모듈이다. 입력 특징맵에 대해 Horizontal Strip Pooling, Vertical Strip Pooling, Channel Attention을 병렬적으로 수행한다.

수평 방향 정보 추출은 식 (1), 수직 방향 정보 추출은 식 (2), 전역 채널 어텐션은 식 (3)과 같이 정의된다.

$$F_h = Conv_{1\times 3}(AdaptiveAvgPool_{1\times w}(X)) \quad (1)$$

$$F_v = Conv_{3\times 1}(AdaptiveAvgPool_{H\times w}(X)) \quad (2)$$

$$F_c = Conv_{1\times 1}(GAP(X)) \quad (3)$$

FAAM은 가로·세로 방향의 연속 구조 정보를 반영하여 Pole 및 Pier와 같이 방향성이 뚜렷한 구조물과 Rain_gutter와 같이 주변 클래스와의 경계가 모호한 소수 클래스의 특징 표현을 보완하기 위해 설계되었다. Horizontal Strip Pooling은 가로 방향으로 길게 분포하는 구조를, Vertical Strip Pooling은 세로 방향으로 길게 분포하는 구조를 포착하기 위한 것이다. Channel Attention은 침수영역 및 구조물 구분에 유효한 특징 채널의 중요도를 재조정한다. 세 분기의 출력은 원래 해상도로 복원된 뒤 결합되고, 1×1 convolution과 sigmoid 연산을 거쳐 Attention 가중치 A를 생성한다. 최종 출력은 식 (4)와 같이 잔차 형태로 정의된다.

$$Y = A \odot X + X \quad (4)$$

여기서 ⊙는 요소별 곱을 의미한다. 해당 구조는 세장형 구조물의 방향성 정보를 명시적으로 보완한다는 점에서 도시홍수 CCTV 장면에 적합하다.

4) Multi-Scale Decoder

Multi-Scale Decoder는 생성된 특징을 단계적으로 융합하여 최종 512×512 세그멘테이션 마스크를 생성한다. 먼저 64×64 Feature Fusion 단계에서는 Detail Branch의 64×64×96 특징, Context Branch의 32×32×160 특징을 1×1 convolution으로 96채널 정렬 후 64×64로 업샘플한 특징, 그리고 Context Branch의 64×64×64 특징을 결합한다. 이를 통해 256채널의 융합 특징을 생성하고, FAAM과 DSConv 블록을 거쳐 64×64×96 특징으로 정제한다. 이후 128×128 Feature Fusion 단계에서는 64×64×96 특징을 128×128로 업샘플한 특징과 Detail Branch의 128×128×64 Skip Feature를 결합하여 160채널의 융합 특징을 생성한다. DSConv 블록을 통해 128×128×64 특징으로 정제하고, 최종 Segmentation Head를 통해 512×512×7 형태의 7-class 세그멘테이션 출력을 생성한다. Auxiliary Head는 학습 단계에서만 사용되며, 64×64×96 특징에 ConvBNReLU와 1×1 convolution을 적용하여 7-class 보조 출력을 생성한다. 보조 출력은 최종 해상도인 512×512로 업샘플되어 손실 계산에 사용된다. 추론 단계에서는 보조 헤드를 사용하지 않으므로, 추가적인 추론 시간은 발생하지 않는다.

5) Training Loss Function

FloodSP-Net의 최종 손실 함수는 Main Loss와 Auxiliary Loss의 가중합으로 식 (5)와 같이 정의된다.

$$L_{total} = L_{CE}(y_{main}, y_{gt}) + \lambda L_{CE}(y_{aux}, y_{gt}) \quad (5)$$

여기서 y_{main}은 최종 출력, y_{aux}는 보조 출력, y_{gt}는 Ground Truth, λ는 Auxiliary Loss 가중치이다. 본 연구에서는 $\lambda = 0.4$로 설정하였다. 두 손실 모두 식 (6)과 같이 역빈도 기반 weighted cross-entropy를 사용하여 클래스 불균형을 보정하였다.

$$L_{CE} = -\sum_{c=1}^{K} w_c y_{gt,c} \log(y_{pred,c}),\ w_c = \frac{1}{R_c + \epsilon} \quad (6)$$

여기서 w_c는 클래스 c의 역빈도 가중치, R_c는 클래스 픽셀 비율, ϵ은 수치 안정성을 위한 상수이다. 이러한 학습 설계는 소수 클래스가 중간 단계와 최종 단계에서 동시에 학습 신호를 받을 수 있도록 한다.

4-2 FloodSP-Net 성능 분석 및 결과

1) 클래스별 성능 분석

표 4는 FloodSP-Net의 클래스별 성능 특성을 보다 구체적으로 분석하기 위해, 선정한 주요 모델 4종과 비교한 결과이다.

FloodSP-Net의 클래스별 IoU 분석 결과, Pole과 Rain_gutter는 각각 0.5562, 0.4785의 IoU를 기록하여 BiSeNetV2 및 Fast-SCNN보다 낮은 성능을 보였다. 이는 두 클래스가 낮은 픽셀 비율과 얇은 형상 특성을 가지므로, 경량화된 네트워크 구조에서 세부 형상 및 경계 정보를 충분히 보존하는 데 한계가 있었기 때문으로 판단된다. 반면 Pier는 0.7291의 IoU를 기록하여 BiSeNetV2의 0.7648보다는 낮았으나, Fast-SCNN 0.6566, PIDNet-S 0.6559, ENet 0.4600보다 높은 탐지 성능을 보였다.

Driveway 클래스에서는 FloodSP-Net이 0.9824의 IoU를 기록하여 BiSeNetV2 0.9792, Fast-SCNN 0.9622, PIDNet-S 0.9450, ENet 0.9120보다 높은 성능을 보였다. Water_area 클래스에서는 0.9487의 IoU를 기록하여 BiSeNetV2 0.9526과 유사한 수준을 보였으며, Fast-SCNN 0.9208, PIDNet-S 0.8666, ENet 0.7877보다 높은 성능을 나타냈다. Water_side 클래스에서는 0.9266의 IoU를 기록하여 BiSeNetV2 0.9365보다는 낮았으나, Fast-SCNN 0.8876, PIDNet-S 0.8601, ENet 0.7707보다 높은 성능을 보였다. 이는 FloodSP-Net이 도로 영역, 침수영역 및 수변부와 같은 주요 클래스에서 비교 경량 모델 대비 높은 탐지 성능을 확보하였음을 의미한다.

종합하면, FloodSP-Net은 Driveway 0.9824, Water_area 0.9487, Water_side 0.9266, Pier 0.7291의 IoU를 기록하여 도로 침수 분석에 필요한 주요 클래스에서 비교적 높은 탐지 성능을 보였다. 다만 Pole과 Rain_gutter에서는 각각 0.5562, 0.4785로 상대적으로 낮은 성능을 보여, 소형·세장형 클래스의 세부 형상 보존에는 추가적인 개선이 필요한 것으로 판단된다.

2) 종합 성능 분석

FloodSP-Net의 성능은 3장에서 사용한 동일한 입력 크기, 옵티마이저, 학습률, 학습률 스케줄러, Epoch, 손실 함수 조건을 기반으로 검증하였다. FloodSP-Net과 선정된 주요 경량 모델인 4종 세그멘테이션 모델의 성능 분석 결과는 표 5와 같다.

표 5. FloodSP-Net과 4종 비교 모델의 성능 비교
Table 5. Performance analysis of FloodSP-Net and 4 models

Model	mIoU	FPS	Parameters
FloodSP-Net	**0.8018**	**556.4**	**1.29M**
BiSeNetV2	0.8594	600.9	2.04M
Fast-SCNN	0.7914	784.5	1.12M
PIDNet-S	0.7560	1459.0	1.67M
ENet	0.6587	404.4	0.21M

먼저 mIoU 측면에서 FloodSP-Net은 BiSeNetV2의 0.8594보다 낮았으나, 파라미터 수 2M 이하의 경량 모델인 Fast-SCNN 0.7914 대비 1.04%, PIDNet-S 0.7560 대비 4.58%, ENet 0.6587 대비 14.31% 높은 성능을 보였다.

FPS 측면에서 FloodSP-Net은 556.4 FPS를 기록하였다. 이는 BiSeNetV2의 600.9 FPS, Fast-SCNN의 784.5 FPS, PIDNet-S의 1459.0 FPS보다 낮지만, ENet의 404.4 FPS보다는 높은 수준이다. FloodSP-Net은 최고 속도 모델은 아니지만, 500 FPS 이상의 추론 속도를 확보하여 실시간 영상 처리가 가능한 경량 모델로 해석할 수 있다.

표 4. FloodSP-Net과 주요 경량 모델 4종의 클래스 IoU 비교
Table 4. Class-wise IoU comparison of FloodSP-Net

Model	Background	Driveway	Pole	Pier	Water_side	Water_area	Rain_gutter
FloodSP-Net	**0.9912**	**0.9824**	**0.5562**	**0.7291**	**0.9266**	**0.9487**	**0.4785**
BiSeNetV2	0.9823	0.9792	0.7724	0.7648	0.9365	0.9526	0.6277
Fast-SCNN	0.9663	0.9622	0.6423	0.6566	0.8876	0.9208	0.5038
PIDNet-S	0.9571	0.9450	0.5978	0.6559	0.8601	0.8666	0.4095
ENet	0.9308	0.9120	0.5054	0.4600	0.7707	0.7877	0.2443

파라미터 수 측면에서 FloodSP-Net은 1.29M으로, 정확도가 높은 BiSeNetV2 2.04M보다 36.76% 적은 파라미터 수를 보였다. 또한 PIDNet-S 1.67M(mIoU 0.7560)보다 적은 파라미터 수에서 더 높은 mIoU를 기록하였다. 유사한 크기의 파라미터 수를 보이는 Fast-SCNN(1.12M, mIoU 0.7914)보다 FloodSP-Net은 비교 우위의 성능을 보였다. 가장 적은 파라미터 수를 기록한 ENet(0.21M, mIoU 0.6587)은 경량성에 있어서 우위가 있다고 판단되나, mIoU가 상대적으로 낮아 정확도 측면에서 한계가 있다고 판단된다.

종합하면, FloodSP-Net은 BiSeNetV2 대비 mIoU와 FPS는 낮지만, 1.29M의 파라미터 수에서 mIoU 0.80 이상과 500 FPS 이상의 추론 속도를 확보하여 실시간 침수영역 탐지에 활용 가능한 모델로 판단된다.

V. 결론 및 향후 연구

본 연구는 도시홍수 CCTV 영상 기반 실시간 침수영역 세그멘테이션 모델의 성능 특성과 한계를 분석하고, 경량 환경에서 적용 가능한 FloodSP-Net을 제안하였다. 부산시 침수위험 복합 데이터셋 89,995장을 활용하여 10종의 세그멘테이션 모델을 동일 조건에서 비교하였으며, mIoU, 클래스별 IoU, FPS, 파라미터 수를 기준으로 모델의 적용 가능성을 평가하였다.

분석 결과, DeepLabV3+ (ResNet-50)는 mIoU 0.9002로 가장 높은 정확도를 보였으나 41.99M의 파라미터 수로 인해 엣지 환경 적용에는 제약이 있었다. BiSeNetV2는 mIoU 0.8594, 600.9 FPS, 2.04M 파라미터로 정확도와 실시간성 측면에서 가장 우수한 균형을 보였다. 반면 Fast-SCNN, PIDNet-S, ENet의 경량 모델은 높은 FPS를 확보하였으나, 상대적으로 낮은 mIoU 성능이 확인되었다.

이에 본 연구에서는 제한된 연산 자원에서 파라미터 수를 줄이면서도 일정 수준의 정확도와 실시간성을 확보하기 위한 FloodSP-Net을 설계하였다. 제안 모델은 1.29M의 파라미터 수에서 mIoU 0.8018, 556.4 FPS를 기록하였다. BiSeNetV2와 비교하면 mIoU와 FPS는 낮았으나, 더 적은 파라미터 수로 0.80 이상의 mIoU와 500 FPS 이상의 추론 속도를 확보하였다. 따라서 FloodSP-Net은 적은 파라미터 수를 기반으로 정확도와 실시간성을 함께 고려한 경량 모델로 판단된다.

한계점으로는 FloodSP-Net이 Pole과 Rain_gutter 클래스에서 각각 0.5562, 0.4785의 IoU를 기록하여 일부 비교 모델보다 낮은 성능을 보였다. 특히 Rain_gutter, Pole 클래스는 픽셀 비율이 낮고 얇은 수직 구조를 가지므로, 경량화된 네트워크 구조에서 세부 형상 정보를 충분히 보존하는 데 한계가 있었던 것으로 판단된다. 또한 본 연구는 부산시 침수위험 복합 데이터셋과 단일 GPU 환경을 기반으로 수행되었으며, 별도의 독립 테스트셋이 아닌 동일 데이터셋 내 검증용 데이터를 성능 평가에 활용하였다는 한계가 있다. 학습 조건 측면에서는 모델 간 비교의 일관성을 확보하기 위해 Epoch을 15로 통일하고 Random Horizontal Flip만을 적용하였으나, 모델별 최적 Epoch 차이와 다양한 시각적 환경 변화를 충분히 반영하지 못한 한계가 있다.

이러한 한계를 극복하기 위해 향후 연구에서는 경계 보존 모듈, 고해상도 특징 융합, 소수 클래스 중심 학습 전략을 적용하여 소형·세장형 구조물의 탐지 성능을 보완하고, 다양한 도시 및 하드웨어 환경에서 모델의 일반화 성능과 실시간 처리 안정성을 추가 검증하고자 한다. 또한 validation loss 기반 조기 종료와 모델별 최적 Epoch 탐색을 통해 각 모델의 최대 성능을 비교할 필요가 있다. 더불어 저조도 시간대와 야간 환경에서 발생하는 밝기 저하 및 경계 흐림 문제를 완화하기 위해 데이터 증강, 전처리 기법, 도시홍수 영상 특화 필터를 결합하고, 영상 조건에 따라 필터 강도를 적응적으로 조절하는 방향으로 연구를 확장하고자 한다. 이를 통해 실시간성과 경량성을 유지하면서도 침수영역 및 주요 구조물의 탐지 성능을 향상시키고자 한다.

감사의 글

본 연구는 한국건설기술연구원 연구운영지원(주요사업) 사업으로 수행되었습니다(20260161001, 홍수 안심도시 실현을 위한 디지털 도시홍수제어 기술 개발).

참고문헌

[1] D. Dodman, B. Hayward, M. Pelling, V. Castan Broto, W. Chow, E. Chu, ... and G. Ziervogel, “Cities, Settlements and Key Infrastructure,” in *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Cambridge, UK: Cambridge University Press, pp. 907-1040, 2022. https://doi.org/10.1017/9781009325844.008

[2] Ministry of the Interior and Safety, 2023 Disaster Yearbook (Natural Disaster), Ministry of the Interior and Safety, Sejong, January 2025.

[3] J. Park, C. Yoon, Y. R. Kim, J.-G. Sung, and M. S. Kim, “A Study on Transportation Response Strategies in Flood Scenarios,” *Journal of Digital Contents Society*, Vol. 26, No. 11, pp. 3251-3257, November 2025. http://dx.doi.org/10.9728/dcs.2025.26.11.3251

[4] AI Hub. Busan City Flood Risk Complex Data [Internet]. Available: https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=71793.

[5] M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. R. Murphy, “FloodNet: A High Resolution

Aerial Imagery Dataset for Post Flood Scene Understanding," *IEEE Access*, Vol. 9, pp. 89644-89654, June 2021. https://doi.org/10.1109/ACCESS.2021.3090981

[6] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, "Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data," *Sensors*, Vol. 19, No. 7, 1486, March 2019. https://doi.org/10.3390/s19071486

[7] Y. Liang, X. Li, B. Tsai, Q. Chen, and N. Jafari, "V-FloodNet: A Video Segmentation System for Urban Flood Detection and Quantification," *Environmental Modelling & Software*, Vol. 160, 105586, February 2023. https://doi.org/10.1016/j.envsoft.2022.105586

[8] Y. Wang, Y. Shen, B. Salahshour, M. Cetin, K. Iftekharuddin, N. Tahvildari, ... J. L. Goodall, "Urban Flood Extent Segmentation and Evaluation from Real-World Surveillance Camera Images Using Deep Convolutional Neural Network," *Environmental Modelling & Software*, Vol. 173, 105939, 2024. https://doi.org/10.1016/j.envsoft.2023.105939

[9] J. Zhao, X. Wang, C. Zhang, J. Hu, J. Wan, L. Cheng, ... and X. Zhu, "Urban Waterlogging Monitoring and Recognition in Low-Light Scenarios Using Surveillance Videos and Deep Learning," *Water*, Vol. 17, No. 5, 707, 2025. https://doi.org/10.3390/w17050707

[10] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," in *Proceeding of the 15th European Conference on Computer Vision - ECCV 2018*, Munich, Germany, pp. 833-851, 2018. https://doi.org/10.1007/978-3-030-01234-2_49

[11] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation," arXiv:1606.02147, 2016. https://doi.org/10.48550/arXiv.1606.02147

[12] R. P. K. Poudel, S. Liwicki, and R. Cipolla, "Fast-SCNN: Fast Semantic Segmentation Network," arXiv:1902.04502, 2019. https://doi.org/10.48550/arXiv.1902.04502

[13] C. Yu, C. Gao, J. Wang, G. Yu, C. Shen, and N. Sang, "BiSeNet V2: Bilateral Network with Guided Aggregation for Real-Time Semantic Segmentation," *International Journal of Computer Vision*, Vol. 129, pp. 3051-3068, 2021. https://doi.org/10.1007/s11263-021-01515-2

[14] H. Pan, Y. Hong, W. Sun, and Y. Jia, "Deep Dual-Resolution Networks for Real-Time and Accurate Semantic Segmentation of Traffic Scenes," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 24, No. 3, pp. 3448-3460, 2023. https://doi.org/10.1109/TITS.2022.3228042

[15] M. Fan, S. Lai, J. Huang, X. Wei, Z. Chai, J. Luo, and X. Wei, "Rethinking BiSeNet for Real-Time Semantic Segmentation," in *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville: TN, pp. 9711-9720, June 2021. https://doi.org/10.1109/CVPR46437.2021.00959

[16] J. Peng, Y. Liu, S. Tang, Y. Hao, L. Chu, G. Chen, ... and Y. Ma, "PP-LiteSeg: A Superior Real-Time Semantic Segmentation Model," arXiv:2204.02681, 2022. https://doi.org/10.48550/arXiv.2204.02681

[17] J. Xu, Z. Xiong, and S. P. Bhattacharyya, "PIDNet: A Real-Time Semantic Segmentation Network Inspired by PID Controllers," in *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, pp. 19529-19539, June 2023. https://doi.org/10.1109/CVPR52729.2023.01871

[18] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers," in *Advances In Neural Information Processing Systems*, Vol. 34, pp. 12077-12090, December 2021.

김민석(Min Seok Kim)

2023년 : 경북대학교 건설방재공학 학사

2023년～현 재: 과학기술연합대학원대학교 박사과정

2023년～현 재: 한국건설기술연구원 도로교통연구본부 UST학생연구원

※관심분야 : 도로안전, 영상처리, 딥러닝

윤천주(Chunjoo Yoon)

2000년 : 인하대학교 자원공학 학사

2002년 : 인하대학교 지리정보공학 석사

2022년 : 서울시립대학교 교통공학 박사

2002년～현 재: 한국건설기술연구원 도로교통연구본부 수석연구원

※관심분야 : 측량, GIS, 도로안전